Knowledge-Based Systems 85 (2015) 80–95

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Computer aided detection of spina bifida using nearest neighbor classification with curvature scale space features of fetal skulls extracted from ultrasound images

Umut Konur [a,*], Fikret S. Gürgen [a], Füsun Varol [b], Lale Akarun [a]

[a] Computer Engineering Department, Boğaziçi University, Istanbul, Turkey
[b] Department of Obstetrics and Gynecology, Medical Faculty of Trakya University, Edirne, Turkey

ARTICLE INFO

Article history:
Received 31 January 2014
Received in revised form 18 April 2015
Accepted 26 April 2015
Available online 5 May 2015

Keywords:
Curvature scale space
Differential turning angle
Spina bifida
Ultrasound
Receiver operating characteristics
Area under the ROC curve

ABSTRACT

This paper addresses the problem of detecting the common neural tube defect of *spina bifida* by a computer aided detection (CAD) system. We propose a method which extracts the curvature scale space (CSS) features of fetal skull contours viewed in the ultrasound (US) modality and performs nearest neighbor (*k*NN) classification on those features having the desired properties of invariance with respect to translation, orientation and scale changes, thus improving robustness. The distance between two sets of CSS features, each set corresponding to the description of the contour of a particular skull, is measured as the cost of matching the two sets of CSS features. Such a CAD system may act as a second observer and help experts in prenatal diagnosis.

Our data possess absolute and relative rarity. The experiments are performed with two different rare class handling methods and over a range of operating conditions. All experiments are based on a group of settings; associated with using either *balanced* or *unbalanced* datasets, employing different types of CSS features and how CSS matching costs are computed. Comparatively evaluating the classification performance of the settings is carried with the aid of the whole-curve metric of area under the receiver operating characteristics (ROC) curve (AUC). Optimal operating conditions for any setting can be identified and some settings reveal advantages over others. The observations indicate that using balanced datasets offers better performance and our proposed version of estimating CSS matching costs is generally superior to the classical method. Furthermore, using enhanced sets of CSS features improves classification accuracy. When classification is performed on balanced data using enhanced CSS features and the matching cost is computed with our proposed technique; one can observe an F-measure of 0.76 along with 70% TP rate (recall), 17% FP rate (false alarms) and 82% precision.

1. Introduction

With the advent of computer science and automatization of many tasks in many fields, realizing computer aided diagnosis (CAD) systems has been an area of interest in biomedicine. CAD usually refers to procedures in medicine to help specialists in the interpretation of medical images for decision-making. Typically, CAD is an interdisciplinary framework consisting of radiological and digital image processing combined with machine learning. Imaging techniques of X-ray, magnetic resonance (MRI), computed tomography (CT) and ultrasound (US) are the main sources of data (i.e. images) to be used as inputs to CAD systems. The structures of interest (i.e. regions of interest (ROI)) in images are usually analyzed to detect the presence of conspicuous structures which help identify particular diseases. CAD systems may only be intended to act as supporting agents and not substitute doctors who are responsible for the final interpretation of images in an absolute and ethical sense. Examples of CAD applications include the diagnosis of various cancer types (e.g. lung, colon, breast) [8,20,28], coronary artery disease [2], mammographic masses [7], peripheral soft tissue masses [6], and more.

In the CAD application subject to this paper, our aim is the automatized detection of the common neural pathology called *spina bifida* (open/split spine) from ultrasound (US) images of fetal skulls acquired in prenatal terms. As the name implies, spina bifida is one of a group of defects known as *neural tube defects* (*NTD*)

* Corresponding author.
E-mail addresses: konur@boun.edu.tr (U. Konur), gurgen@boun.edu.tr (F.S. Gürgen), fgvarol@yahoo.com (F. Varol), akarun@boun.edu.tr (L. Akarun).

http://dx.doi.org/10.1016/j.knosys.2015.04.021
0950-7051/

related to the spine and spinal cord. In neural development of embryos, a tissue called the neural plate folds and forms a tube, which then folds into the spinal cord. Incorrect folding of the neural plate causes the spina bifida defect [9], whose result is an abnormally formed section of the spinal column. Abnormality occurs at some vertebral column location, referred to as the lesion level. People suffering from spina bifida experience loss of body control below the lesion. The higher the lesion level (the more cranial or the closer to the brain), the more severe the defect. Loss of body control may appear as problems in bladder control, sensation loss and paralysis. Fig. 1 shows the sagittal section of a defective spine of a fetus viewed with US. The prevalence of spina bifida is 1–2 cases per 1000 births worldwide and the incidence is observed to vary up to 3–4 cases per 1000 in some populations. Although what causes the injury is not well known; the consumption of folic acid, a type of vitamin B, by pregnant women shows to prevent up to 70% of neural tube defects including spina bifida.

US examination is a convenient tool to discover neural tube defects in the prenatal stage. Detection of the defect before birth leads of careful planning and effective remedy. Surgical treatment after birth and fetal surgery may be possible, however, most pregnancies with neural tube defects are terminated because of the poor future quality of life of newborns having such defects. Observing the spine itself for detection is not a necessity because fetal heads contain markers indicating the presence of spina bifida. The so-called *lemon sign* [22] that appears when the frontal bones of the skull look flattened and inwardly bent, is a very typical marker. Fig. 2 shows the transcerebellar section of a malformed skull of a defective fetus with lemon sign.

Automatically deciding whether a fetus is associated with the spina bifida defect (i.e. label 1) or not (i.e. label 0) is obviously a classification problem. Robust classification, from a general viewpoint, involves a number of subtasks such as preprocessing inputs to represent them by means of descriptive features and treating the extracted features with appropriate machine learning methods to assign *patterns/classes/labels* to them with objectives of optimizing pre-defined performance criteria. Each subtask may further be composed of other subtasks. Specific problems; varying in aspects of input and feature modalities, the number of samples available for learning, the distribution of samples belonging to different classes, the primary objective of classification and how success is perceived; require specific actions to be undertaken. The overall success depends on the selection of algorithms used in each subtask with careful consideration of each detail. Assessing classification performance and generalizing outcomes is also a matter of data sets and methodology employed throughout experiments [10–12,30,31].

Since spina bifida and lemon sign are tightly-coupled and the latter is associated with outer lines of skulls, features extracted from skull contours are supposed to be appropriate for classification. Features along a contour can be captured at sampled points

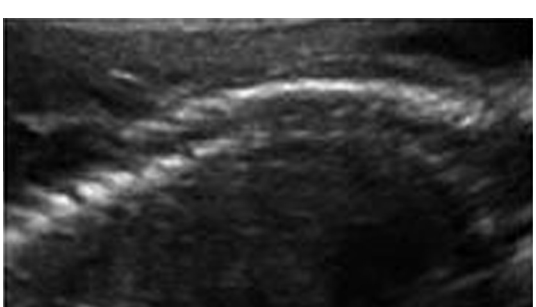

Fig. 1. Sagittal spine of a fetus with spina bifida.

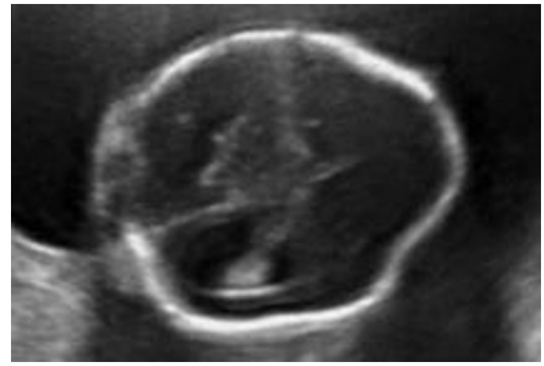

Fig. 2. Transcerebellar fetal head with lemon sign.

and their combination at all points in a well-enough resolution provide a description of the whole contour. The naturally-arising concept is *curvature* of points that characterizes how the contour (or curve) bends at the corresponding points. In the curvature scale space (CSS) representation of curves [26,27], the description is performed by features invariant under translation, rotation and scale. The properties of invariance are definitely very valuable in robust classification because, as applicable to the classification of skull contours, identical shapes must output identical classification results no matter how they are translated, oriented and scaled. One property of the CSS representation is that features for different contours may be of different sizes and classifiers that work on the principle of constructed models using feature vectors of equal size are impractical. In the availability of a scheme that can estimate how similar (or different) two curves are, the lazy k-nearest neighbor (kNN) can be used to compare curves to others with known labels and decide based on similarity. Fortunately, a matching procedure that measures the distance between CSS features of two curves [1] was designed justifying the use of nearest neighbor classifiers.

In the scheme that we propose in this paper, contour lines of fetal skulls are presented to the system and invariant features are computed at multiple scales. The CSS representation [1,24–27] of contours produces a map of sampled points with curvatures of zero value at a number of scales (levels of detail (LOD)). As a result, the CSS image of the associated contour from which features can be extracted is obtained. Parametric representations for curves allow analytic solutions in curvature computation. With curves in digital images, such functional forms are not available and techniques exploiting concepts related to the definition of curvature can be utilized. We use differential turning angles (dTA) [16–19] and their scalograms (dTASS – differential turning angle scale space) to obtain points of zero curvature (zero-crossings) at all scales of consideration. Different scales of a contour correspond to its smoothed versions, each with a different standard deviation σ of a Gaussian kernel used for smoothing. CSS images are fed to the CSS matching algorithm of Abbasi et al. [1] in a pairwise fashion to output a matching (similarity or distance) score for the two CSS images (hence for the two contours). Classification is performed by kNN.

Fig. 3 shows the block diagram of the proposed CAD system which also includes the module for segmenting input US images to isolate fetal skulls from their surroundings. The work of this article assumes that segmentation, which is a challenging problem on its own right, has been solved and skull shapes (i.e. contours) are ready to be processed by the other modules of the system. Figs. 4 and 5 display four typical US images of fetal skulls viewed as transcerebellar sections and their associated shapes (contours), respectively.